HAMBURG 2016

Multi-dimensional analysis of HRTF using tensor singular value decomposition

Daehyuk SON[1]; Youngjin PARK[2]

[1] Korea Advanced Institute of Science and Technology, South Korea

[2] Korea Advanced Institute of Science and Technology, South Korea

ABSTRACT

Researches on virtual three-dimensional sound have been implemented actively to boost a sense of reality to the listener for various multimedia contents. Head-related transfer function (HRTF), an acoustic transfer function between pressure of sound source and listener's ear drum, plays an important role in virtual sound technology. HRTFs of a random subject should be measured with different angles and distances of sound sources to produce virtual sound at any desired position for the listener. However, it is hard to measure all points of three-dimensional sound sources, so many HRTF personalization or customization techniques based on HRTF database have been developed. Previous researches mainly focused on median or horizontal plane only which is two dimensional analysis. In this paper, we analyze HRTF in multi-dimensional point of view using tensor singular value decomposition method because this method can handle many dimensions. Time (or frequency), azimuth, elevation, and subject dimension were adopted as controlling axes. CIPIC database was used and analyzed results were compared with ordinary analyzing method such as principal component analysis.

Keywords: Three-dimensional sound, head-related transfer function, multi-dimensional analysis, I-INCE Classification of Subjects Number(s): 74.8

1. INTRODUCTION

Virtual three-dimensional sound technology represents a system or technology that makes a listener feels like virtual sound located at any random position in three-dimensional space by using headphones or speakers. It is known that head-related transfer function, so called HRTF, plays a key role to generate and reproduce virtual sound because HRTF contains information caused by diffraction and difference due to physical shape of the human body. HRTF is defined as an acoustic transfer function between sound pressure of sound source and that measured in front of listener's ear drum (1, 2). HRTF usually represents frequency domain transfer function and in time domain, head-related impulse response (HRIR) is defined as impulse response function.

Because HRTF contains spatial information of sound source, useful HRTF usually measured with changing azimuth, elevation, and distance of a lot of sound source position. Therefore, tons of dataset is necessary to implement virtual sound with high-fidelity for particular listener. To overcome this limitation a number of research groups developed their own HRTF measuring systems and HRTF databases to make people utilize HRTF without measuring specific listener (3, 4, 5). The related researches are the methods for modeling, personalization and interpolation of HRTF or HRIR. With previous researches, many characteristics of HRTF have been revealed based on a lot of different approach (6, 7, 8).

Although many methods for analyzing HRTF has been developed, but this work mainly focused on observe HRTF dataset in the view of multi-dimensional subspace. Subspace composed of frequency, azimuth, elevation and subject axis is used and HRTF data is obtained from well-known CIPIC HRTF database. Multi-dimensional analysis is started from the idea; projection of

[1] infinitude@kaist.ac.kr

[2] yjpark@kaist.ac.kr

multi-dimensional dataset in smaller-dimensional subspace. By using the method named tensor singular value decomposition, multi-dimensional subspaces are observed especially core tensor and scale factor of those spaces.

2. Research backgrounds

2.1 Head related transfer function

Head-related transfer function (HRTF) is a response that characterizes how an ear receives a sound from a point in space; a pair of HRTFs for two ears can be used to synthesize a binaural sound that seems to come from a particular point in space (9). More specifically, the HRTF is defined as the sound pressure at the eardrum (P_s) divided by the sound pressure measured at the position of the head center with the head absent (P_{ff}). This definition is represented in Figure 1. In general, the HRTF can be obtained from the theoretical computation and from the experimental measurement in an anechoic chamber.

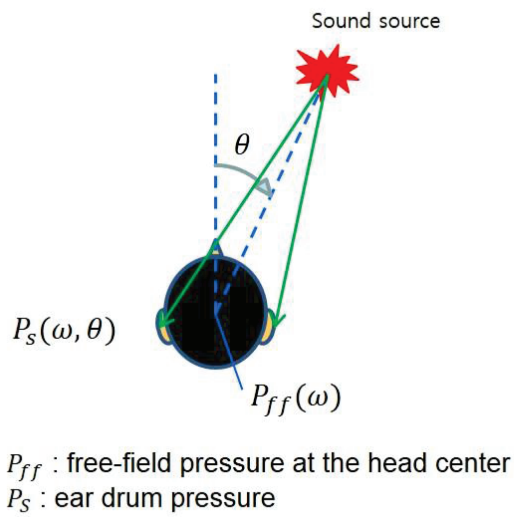

Figure 1 – Head-related transfer function

2.2 Tensor singular value decomposition

Tensor singular value decomposition (Tensor-SVD) method is an expansion method of ordinary matrix singular value decomposition. The word tensor represents a higher-order data, for instance, a matrix is a second-order tensor and a vector is a first-order tensor. With tensor-SVD method, huge data can be controlled and observed because of its characteristics. As an example, third-order tensor can be decomposed by this method like Figure 2 (10).

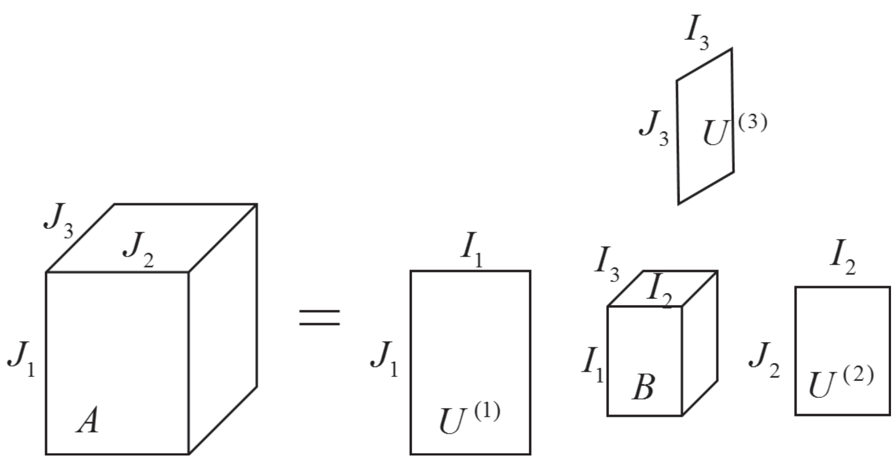

Figure 2 – Diagram of tensor-SVD for decomposing third-order tensor A

Tensor A can be decomposed with core tensor B and factors for each axis $U^{(1)}$, $U^{(2)}$, and $U^{(3)}$. This decomposition process can be represented as an equation (1). Multiplication of B with $U^{(1)}$ means that every column of B has to be multiplied from the left with $U^{(1)}$. Almost all properties of tensor-SVD can be regarded as extension form of matrix SVD.

$$A = B \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \qquad (1)$$

3. Multi-dimensional analysis of HRTF

3.1 Dimension definition

Before mentioning multi-dimensional analysis, dimensions of HRTF is necessary to be defined. In CIPIC HRTF database, there are 45 subjects' HRIRs in time domain. Because impulse responses contain interaural time difference information, analysis is accomplished in frequency domain to prevent result caused by these differences. Especially, we focused on magnitude spectra of HRTFs so first dimension is selected as frequency dimension (F). After that, azimuth (A) and elevation (E) angle of sound source is chosen as second and third dimension to observe characteristics of different incidence angles. Lastly, subject (S) axis is appointed as fourth dimension to observe variation caused by differences among subjects. In short, we analyzed fourth-order tensor which is composed of frequency, azimuth, elevation and subject as equation (2). In this equation, the word C represents core tensor calculated by decomposition process.

$$HRTFs = C \times_F U^{(F)} \times_A U^{(A)} \times_E U^{(E)} \times_S U^{(S)} \qquad (2)$$

Analysis is implemented in each dimension separately in following contents.

3.2 Elevation dimension analysis

Firstly, elevation dimension is controlled and observed in frequency domain. This analysis basically have a purpose to observe characteristics generated in elevation axis. For the analysis, HRTFs in CIPIC database are aligned in fourth-order tensor as represented in equation (2). After that, elevation dimension is compressed from 49 to 10. Modeling accuracy was calculated about 96% which means decomposed data can express original dataset with quite high accuracy.

We focused on scale factor for elevation axis and core tensor of decomposition. In Figure 3, scale factors from first to fourth in elevation dimension are represented.

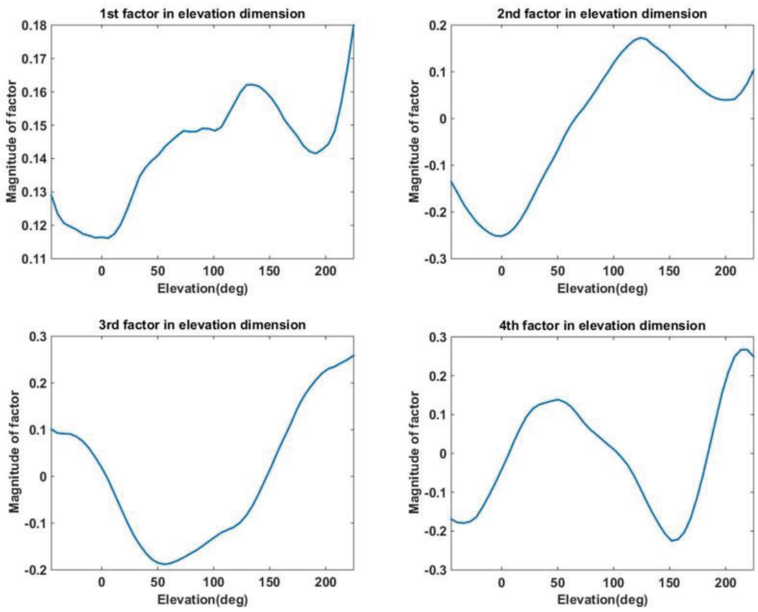

Figure 3 – Scale factors from first to fourth in elevation dimension

In Figure 3, scale factors in elevation dimension show tendency relevant to changing elevation angle. Theoretically, first scale factor has greater variance than other scale factors when decomposition process accomplished. Therefore, first order of core tensor which can be regarded as first basis function is observed like Figure 4. Although first order of core tensor is composed of frequency, azimuth and subject dimension, but variation from subject dimension is approximated by averaging core tensor for subject axis. It means first basis function is expressed in azimuth-frequency domain.

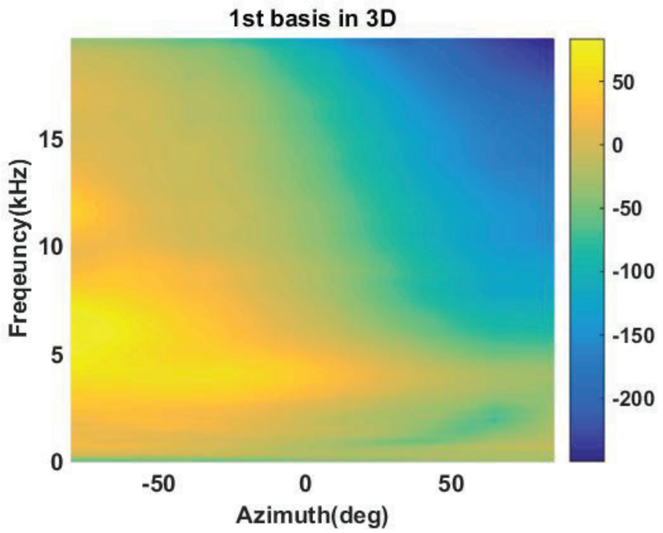

Figure 4 – First basis function in azimuth-frequency domain

First basis shows peaks for left side HRTFs and notches for right side HRTFs. This is because, we used HRTFs measured for left ear. In future, there exists possibilities to personalize or generalize HRTFs by modeling these basis functions and scale factors.

3.3 Azimuth dimension analysis

In this section, azimuth dimension is controlled and observed in frequency domain. This analysis basically have a purpose to observe characteristics generated in azimuth axis. For the analysis, HRTFs in CIPIC database are aligned with same method in previous section. After that, azimuth dimension is compressed from 25 to 10. Modeling accuracy was calculated about 97% which means decomposed data can represent original dataset with quite high accuracy.

Also, we focused on scale factor for azimuth axis and core tensor of decomposition. In Figure 5, scale factors from first to fourth in azimuth dimension are expressed.

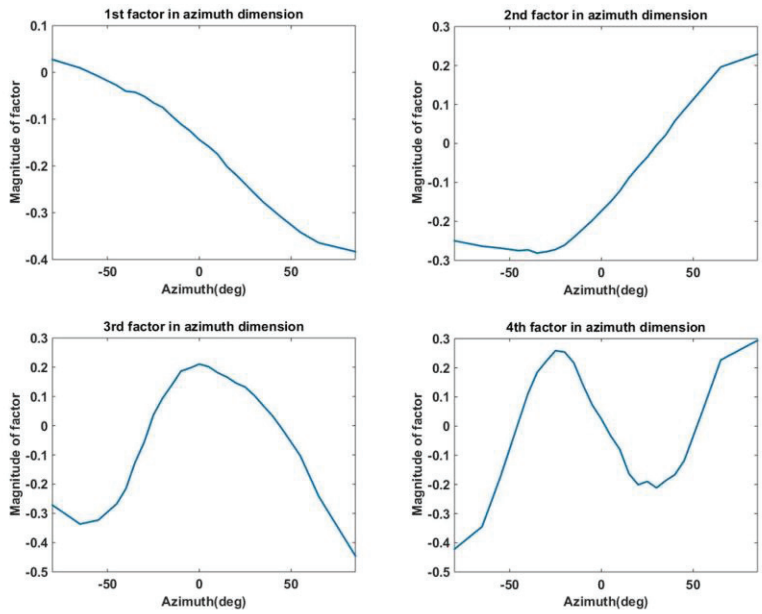

Figure 5 – Scale factors from first to fourth in azimuth dimension

In Figure 5, scale factors in azimuth dimension show tendency relevant to changing azimuth angle. Especially, first and second scale factors represent monotonic relation with azimuth angle which can be easily modeled. Theoretically, first scale factor has greater variance than other scale factors when decomposition process accomplished. Therefore, first order of core tensor which can be regarded as first basis function is observed like Figure 6. Although first order of core tensor is composed of frequency, elevation and subject dimension, but variation from subject dimension is approximated by averaging core tensor for subject axis. It means first basis function is expressed in elevation-frequency domain.

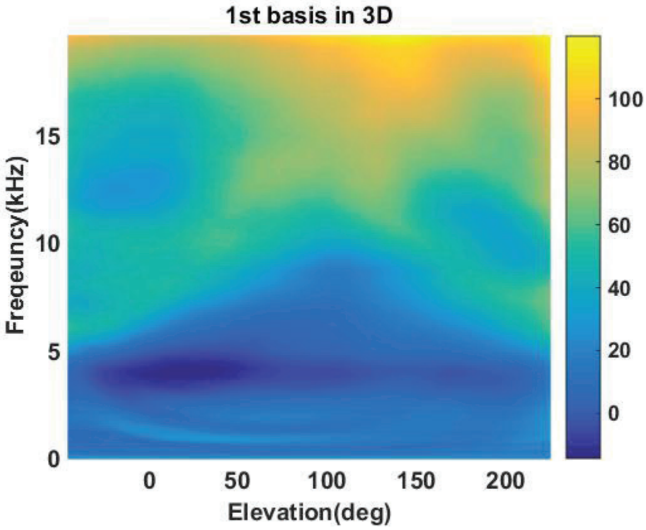

Figure 6 – First basis function in elevation-frequency domain

First basis shows peaks and notches pattern which are similar to HRTFs in median plane. Characteristics such as first pinna notch, shoulder and torso effect can be observed. In future, there exists possibilities to personalize or generalize HRTFs by modeling these basis functions and scale factors

4. CONCLUSIONS

Multi-dimensional analysis of HRTF based on tensor singular value decomposition is proposed in this paper. To control higher order data and observe characteristics for each order, we used tensor-SVD method as analyzing tool. In frequency domain, four kinds of dimensions are selected and data was aligned; frequency, azimuth, elevation, and subject. Observation and analysis is implemented in directional dimensions which are azimuth and elevation. Also, first basis from core tensor and scale factors obtained by each dimension are analyzed. For both dimension, scale factors showed close relation with that dimension that means these factors can be modeled or designed with various approach. First basis function showed features similar to median plane or horizontal plane in azimuth and elevation analysis. In future, different analyzing approach using this method and personalization or generalization of HRTF will be accomplished.

ACKNOWLEDGEMENTS

This work was supported by the Brain Korea 21 Plus Project.

REFERENCES

1. J. Blauert, Spatial Hearing : The Psychophysics of Human Sound Localization, MIT Press, 1983.
2. D. J. Kistler and F. L. Wightman, A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction, J. Acoust. Soc. Am. 1992:91(3):1637-1674.
3. V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, "The CIPIC HRTF database", IEEE Workshop on Applications of Signal Processing to Audio and Acoustics 2001, 2001:99-102.

4. http://recherche.ircam.fr/equipes/salles/listen/
5. H. Moller, M. F. Sorensen, D. Hammershoi and C. B. Jensen, "Head-related transfer functions of human subjects," J. Audio Eng. Soc. 1995:43(5):300-321.
6. P. Majdak, M. J. Goupell and B. Laback. 3-D localization of virtual sound sources: effects of visual environment, pointing method, and training. Atten Percept Psychophys. 2010;72(2):454-469.
7. V. Larcher et al., "Study and Comparision of efficient methods for 3-D audio spatialization based on linear decomposition of HRTF data", Audio Engineering Society Convention, 2000: 5097.
8. M. Rothbucher, H. Shen, and K. Diepold, "Dimensionality reduction in HRTF by using multiway array analysis", Human Centered Robot Systems Cognitive Systems Mongraphs, 2009:6:103-110.
9. https://en.wikipedia.org/wiki/Head-related_transfer_function
10. L. D. Lathauwer, B. D. Moor and J. Vandewalle, "A Multilinear singular value decomposition", SIAM J. Matrix Anal. Appl., 2000:21:1253-1278.